LA-UR -81-3546

TITLE: A COMPARISON OF TESTS FOR SPATIALLY DISTRIBUTED MATERIALS BALANCES

MASTER

AUTHOR(S): R. S. Picard, A. S. Goldman, J. P. Shipley

SUBMITTED TO: 1981 DOE Statistical Symposium Proceedings

— DISCLAIMER —

University of California

LOS ALAMOS SCIENTIFIC LABORATORY
Post Office Box 1663 Los Alamos, New Mexico 87545
An Affirmative Action/Equal Opportunity Employer

Form No. 836 R3
St. No. 2629
12/78

UNITED STATES
DEPARTMENT OF ENERGY
CONTRACT W-7405-ENG. 36

A COMPARISON OF TESTS FOR SPATIALLY DISTRIBUTED MATERIALS BALANCES
PART I

R. Picard, A. Goldman, and J. P. Shipley

Los Alamos National Laboratory
Los Alamos, New Mexico 87545

Abstract

The two-box case illustrates aspects of a one-sided test of hypothesis of means of a bivariate normal with known correlation matrix. One important application is testing for materials loss at a nuclear processing facility. Comparisons among five test procedures reveal that a simple sum approach is quite competitive with more complex methods.

I. Introduction

Facilities engaged in processing nuclear materials are typically composed of control units or materials balance areas where nuclear matter is measured. An idealized system with two materials balances is depicted in Fig. 1. If all material is accounted for,

$$\text{Input} - \text{Waste Stream 1} - \text{Waste Stream 2} - \text{Output} = 0 \ .$$

If loss or diversion of material took place in the first "box," then the first balance would satisfy

$$x_1 = \text{Input} - \text{Waste Stream 1} - \text{Transfer} > 0 \ .$$

The objective of safeguards work is "the timely detection of diversion of significant quantities of nuclear material... ."[1] Independent measurements of known precision are obtained to monitor the process, and the problems of interest are detection (is all material accounted for?), and location (if some material is missing, from which box(es) was it diverted?).

Mathematical formulation of these problems is as follows. Let x_i denote the i^{th} observed materials balance. Because adjacent boxes depend on the same transfer measurement, the $\{x_i\}$ are negatively correlated. Let μ_i denote the i^{th} actual materials balance, and $\underline{\mu}$ the vector of $\{\mu_i\}$. The hypothesis of interest for the detection problem is that no loss (diversion) has taken place in any of the boxes, that is,

$$H_0: \underline{\mu} = \underline{0} \ .$$

If loss is detected, the location problem entails identification of which of the $\{\mu_i\}$ are positive.

Of course, the above is an oversimplified version of what is encountered in practice. Typically, systems are composed of 5-10 boxes, each box having a number of associated transfers, inventories, and waste streams. Difficulties with material holdup and recalibration of measurement devices must be dealt with. Also, problems can be sequential as well as spatial in nature. Detection of losses in a given box over time is often another important issue and has received some attention (Cobb[2]). The groundwork provided here with the two-box problem should lead to valuable procedures for the more genera

II. Methods for Detection

The two-box problem reduces to testing $H_0: \underline{\mu} = \underline{0}$ against the one-sided alternative

$$H_A: \underline{\mu} \geq \underline{0} \ , \ \max\{\mu_1, \mu_2\} > 0 \ .$$

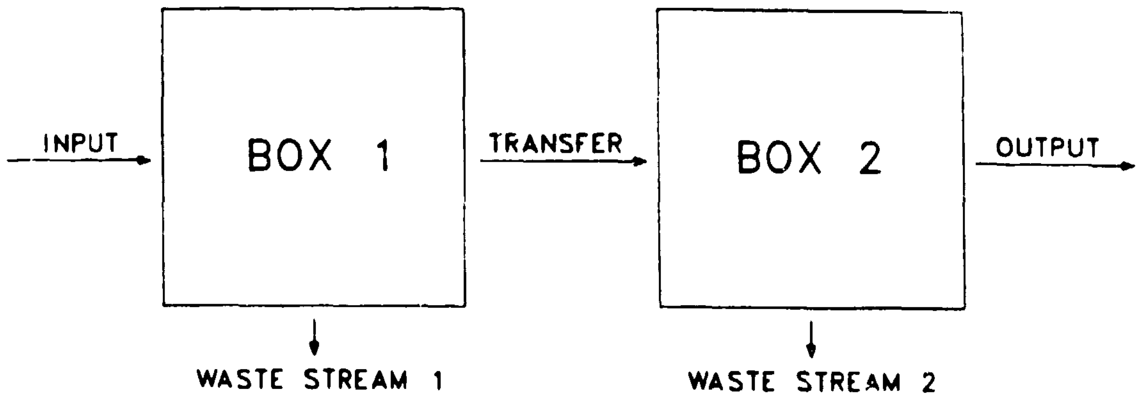

Fig. 1. An example of a materials control unit consisting of two materials balance areas.

Measurement errors tend to be closely approximated by a normal distribution, and here it is assumed that the observed materials balances satisfy

$$x \sim N(\mu, \Sigma) \quad \text{for } \Sigma = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \text{ and } \rho < 0 \text{ known} \quad .$$

A variety of procedures have been proposed for such testing. Two that are quite simple and mimic common quality control techniques are described below.

A. <u>The SUM Test</u>. Somewhat akin to the CUSUM statistic, balances x_1 and x_2 are summed and used to test for diversion. The null hypothesis is rejected at level α when

$$(x_1 + x_2)/\sqrt{2 + 2\rho} > z_{1-\alpha} \quad ,$$

where $z_{1-\alpha}$ denotes the $100(1 - \alpha)$ percentile of the standard normal distribution.

B. <u>The MAX Test</u>. In addition to examining the SUM, balances can be checked individually, similar to a Shewart approach. If any of the tests is significant, H_0 is rejected. The critical region is

$$\max\{x_1, x_2, (x_1 + x_2)/\sqrt{2 + 2\rho}\} > z^*_{\rho,1-\alpha} \quad ,$$

where $z^*_{\rho,1-\alpha}$ is chosen so that the test has size α.

A third procedure to consider is the standard multivariate treatment for the unrestricted alternative $H_A: \mu \neq 0$. Specifically, this is as follows.

C. <u>A χ^2 Test</u>. The statistic

$$\chi^2 = x'\Sigma^{-1}x$$

has a χ^2 distribution with two degrees of freedom under H_0. The test rejects when χ^2 exceeds the $100(1 - \alpha)$ percentile of a χ^2_2.

A desirable quality of these procedures is the simplicity of computations. More complex tests exist in the area of one-sided alternatives in the multidimensional context. Two are given here.

D. <u>The Likelihood Ratio Test (LRT)</u>. This test rejects for large values of

$$\bar{\chi}^2 = x'\Sigma^{-1}x - (x - \hat{\mu})'\Sigma^{-1}(x - \hat{\mu})$$

where $\hat{\underset{\sim}{\mu}}$ is the maximum likelihood estimator of $\underset{\sim}{\mu}$ under H_A. Critical values are determined from the relation (Kudo[3])

$$P\{Y^2 \leq c^2\} = \Phi(c) - \{(\pi - \cos^{-1}\rho)/2\pi\}e^{-c^2/2} ,$$

for $c > 0$ and Φ the standard normal distribution function.

E. A Conditional Probability Test (CPT). Conditional on the observed x, Shirahata[4] develops the CPT by defining the sets

$$A(\underset{\sim}{x}) = \left\{ \underset{\sim}{y} : \frac{f_0(x)}{f_\mu(x)} < \frac{f_0(y)}{f_\mu(y)} \quad \text{for all } \underset{\sim}{\mu} \in H_A \right\}$$

and

$$B(x) = \left\{ \underset{\sim}{y} : \frac{f_0(x)}{f_\mu(x)} > \frac{f_0(y)}{f_\mu(y)} \quad \text{for all } \underset{\sim}{\mu} \in H_A \right\} ,$$

where $f_\mu(x)$ denotes the probability density function of a bivariate normal with mean μ and covariance Σ. Loosely speaking, $A(x)$ contains points that are more supportive of H_0 than x, whereas $B(x)$ contains points that are less supportive. H_0 is rejected for small ratios of the probabilities

$$\int_{A(\underset{\sim}{x})} f_0(\underset{\sim}{y})\, d\underset{\sim}{y} \Big/ \int_{B(\underset{\sim}{x})} f_0(\underset{\sim}{y})\, d\underset{\sim}{y} .$$

This test is practical only for negative ρ, as both $A(x)$ and $B(x)$ have small probability under H_0 when ρ is positive.

The LRT and CPT are relatively simple computations in the two-box case, but become complex when the problem is extended to four or more boxes. Solution for $\hat{\mu}$ in the LRT requires quadratic programming for nonzero ρ. For the CPT, determination of the sets $A(\underset{\sim}{x})$ and $B(\underset{\sim}{x})$ together with their associated probabilities is not easy. Critical values for either of the test statistics must be obtained numerically.

Other procedures for the one-sided problem that have been proposed, such as the test of Hafsma and Smid,[5] are not considered here because they are not widely advocated.

III. Summary of Comparisons

Though much work has appeared concerning the statistical theory behind one-sided testing (Perlman;[6] Eaton[7]), no comprehensive comparison of the various approaches appears to have been made. To this end, simulation work was undertaken and some results are displayed in Tables I-V and Figs. 2-4.

Perhaps most surprising was the strong performance of the SUM test, which has been almost completely ignored in the literature. This test proved superior to all others in finding diversion on the equiangular line, where the diverted quantity is equally divided among the boxes. It is here that detection of the quantity is generally least likely; thus, the procedure works well against the "optimal" diversion strategy. It should be noted that lack of exact measurement often prevents a divertor from attaining this strategy, and alternatives "close" to the equiangular line become important. Here too the SUM has a narrow advantage over the LRT and CPT, though this is offset by larger disadvantages when loss is confined to one box. However, ease of application is an important consideration in many practical circumstances so that the slight overall loss of power may be acceptable.

The standard χ^2 test fares poorly, confirming long-held contentions by Nüesch[8] and others that the usual approach to unrestricted alternatives does not have good properties in the one-sided case. Power of the MAX test is nearly identical to the LRT for ρ near zero, but the procedure is less effective by comparison as the strength of the correlation increases. Shirahata's[4] claim that the CPT is more powerful than the LRT appears unsubstantiated. Results here are in accordance with previous conclusions that CPT has a slight edge over LRT near the equiangular line; nonetheless, the reverse holds near the axes--a point not mentioned to date.

In short, for the two-box problem, the LRT and CPT share the best overall power properties, whereas the SUM approach is quite competitive and very easy to implement in a practical setting. The MAX and χ^2 tests can be equally effective in isolated situations.

References

1. "The Structure and Content of Agreements Between the Agency and States Required in Connection with the Treaty on the Non-Proliferation of Nuclear Weapons," INFCIRC/153 (1972), p. 9.

2. D. D. Cobb, "Sequential Tests for Near-Real-Time Accounting," Nucl. Mater. Manage. X, 62-70 (198[illegible]).

3. A. Kudo, "A Multivariate Analogue of the One-Sided Test," Biometrika 50, 403-418 (196[illegible]).

4. S. Shirahata, "An Approach to a One-Sided Test in the Bivariate Normal Distribution," Biometrika 65, 61-67 (1978).

5. W. Schaafsma and L. J. Smid, "Most Stringent Somewhere Most Powerful Tests Against Alternatives Restricted by a Number of Linear Inequalities," Ann. Math. Stat. 37 1161-1172 (1966).

6. M. D. Perlman, "One-Sided Testing Problems in Multivariate Analysis," Ann. Math. Stat. 40, 549-567 (1969).

7. M. L. Eaton, "A Complete Class Theorem for Multidimensional One-Sided Alternatives," Ann. Math. Stat. 41, 1884-1888 (1970).

8. P. Nüesch, "On the Problem of Testing Location in Multivariate Problems for Restricted Alternatives," Ann. Math. Stat. 37, 113-119 (1966).

TABLE I

SUM STATISTIC
(Table entries are probabilities of detection ($\alpha = 0.05$))

Total Diversion
Units are multiples of the (common) standard deviation of each box.

	0.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0
$\rho = -0.15$	0.1038	0.1882	0.3057	0.4487	0.5996	0.7387	0.8412	0.9187
$\rho = -0.5$	0.1248	0.2534	0.4362	0.6315	0.7[illegible]85	0.9064	0.9697	0.9909
$\rho = -0.85$	0.2265	0.[illegible]643	0.8573	0.9785	0.9980	0.9998	1	1

Notes: For the SUM statistic, probability of detection is independent of diversion strategy. Also, these simulated values are to be used for comparison with the other procedures; slightly more accurate values can be obtained through use of tables of the normal distribution.

TABLE II

LIKELIHOOD RATIO TEST (LRT)

(Table entries are probabilities of detection (α = 0.05))

	Diversion Ratio Box 1:Box 2	Total Diversion: Units are multiples of the (common) standard deviation of each box.							
		0.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0
ρ = -0.15	3:3	0.0957	0.1683	0.2699	0.4052	0.5423	0.6829	0.8022	0.8852
	3:2	0.0954	0.1687	0.2750	0.4105	0.5524	0.6934	0.8133	0.8942
	3:1	0.0973	0.1769	0.2935	0.4375	0.592[illegible]	0.7396	0.8523	0.9245
	3:0	0.1016	0.2025	0.3555	0.5403	0.7239	0.8601	0.9412	0.9803
ρ = -0.5	3:3	0.1174	0.2350	0.4107	0.5998	0.7679	0.8866	0.9578	0.9873
	3:2	0.1173	0.2373	0.4125	0.6007	0.7731	0.8891	0.9593	0.9878
	3:1	0.1177	0.2439	0.4252	0.6180	0.7925	0.9056	0.9668	0.9906
	3:0	0.1223	0.2597	0.4619	0.6776	0.8440	0.9421	0.9845	0.9967
ρ = -0.85	3:3	0.2231	0.5591	0.8535	0.9764	0.9971	0.9988	1	1
	3:2	0.2232	0.5605	0.8538	0.9773	0.9974	0.9999	1	1
	3:1	0.2235	0.5632	0.8553	0.9790	0.9977	0.9999	1	1
	3:0	0.2241	0.5690	0.8633	0.9806	0.9987	0.9999	1	1

TABLE III

CONDITIONAL PROBABILITY TEST (CPT)

(Table entries are probabilities of detection (α = 0.05))

	Diversion Ratio Box 1:Box 2	Total Diversion: Units are multiples of the (common) standard deviation of each box.							
		0.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0
ρ = -0.15	3:3	0.0988	0.1852	0.2981	0.4450	0.5902	0.7302	0.8367	0.9131
	3:2	0.0987	0.1856	0.2982	0.4459	0.5918	0.7338	0.8398	0.9171
	3:1	0.0992	0.1875	0.3046	0.4536	0.6097	0.7519	0.8563	0.9292
	3:0	0.1014	0.1932	0.3265	0.4874	0.6602	0.8020	0.9006	0.9602
ρ = -0.5	3:3	0.1170	0.2405	0.4200	0.6148	0.7867	0.8986	0.9656	0.9890
	3:2	0.1166	0.2419	0.4219	0.6141	0.7869	0.8986	0.9662	0.9895
	3:1	0.11[illegible]	0.24[illegible]4	0.4283	0.6221	0.7947	0.9068	0.9678	0.9913
	3:0	0.1194	0.2513	0.4439	0.6502	0.8211	0.9281	0.9793	0.9953
ρ = -0.85	3:3	0.2198	0.5591	0.8529	0.9780	0.9977	0.9998	1	1
	3:2	0.2199	0.5605	0.8525	0.9773	0.9975	0.9999	1	1
	3:1	0.2194	0.5632	0.8545	0.9790	0.9977	0.9999	1	1
	3:0	0.2206	0.5665	0.8596	0.9798	0.9987	0.9999	1	1

TABLE IV

THE MAX TEST

(Table entries are probabilities of detection (α = 0.05)

	Diversion Ratio Box 1:Box 2	Total Diversion Units are multiples of the (common) standard deviation of each box.							
		0.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0
ρ = -0.15	3:3	0.0948	0.1610	0.2583	0.3858	0.5239	0.6655	0.7863	0.8721
	3:2	0.0951	0.1624	0.2664	0.3900	0.5299	0.6733	0.7957	0.8621
	3:1	0.0973	0.1733	0.2858	0.4251	0.5785	0.7269	0.8438	0.9153
	3:0	0.1017	0.2028	0.3577	0.5438	0.7267	0.8609	0.9472	0.9804
ρ = -0.5	3:3	0.1054	0.2031	0.3489	0.5294	0.7077	0.8457	0.9334	0.9803
	3:2	0.1067	0.2030	0.3529	0.5340	0.7139	0.8486	0.9360	0.9807
	3:1	0.1076	0.2111	0.3684	0.5568	0.7374	0.8709	0.9485	0.9846
	3:0	0.1130	0.2391	0.4336	0.6384	0.8175	0.9235	0.9766	0.9949
ρ = -0.85	3:3	0.1617	0.4303	0.7566	0.9432	0.9928	0.9995	1	1
	3:2	0.1628	0.4309	0.7574	0.9443	0.9928	0.9995	1	1
	3:1	0.1650	0.4348	0.7613	0.9457	0.9934	0.9996	1	1
	3:0	0.1688	0.4555	0.7823	0.9535	0.9953	0.9997	1	1

TABLE V

THE χ^2 TEST

(Table entries are probabilities of detection (α = 0.05)

	Diversion Ratio Box 1:Box 2	Total Diversion Units are multiples of the (common) standard deviation of each box.							
		0.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0
ρ = -0.15	3:3	0.0607	0.0952	0.1591	0.2536	0.3800	0.5210	0.6627	0.7860
	3:2	0.0609	0.0982	0.1629	0.2622	0.3909	0.5319	0.6[illegible]65	0.7971
	3:1	0.0625	0.1047	0.18[illegible]8	0.3004	0.4400	0.5935	0.7392	0.8520
	3:0	0.0673	0.1328	0.2525	0.4230	0.6091	0.7783	0.8927	0.9566
ρ = -0.5	3:3	0.0679	0.1296	0.2424	0.4104	0.5973	0.7653	0.8840	0.9565
	3:2	0.0680	0.1305	0.2450	0.4169	0.6025	0.7719	0.8882	0.9585
	3:1	0.0706	0.[illegible]370	0.2627	0.4410	0.6329	0.8009	0.9098	0.9682
	3:0	0.0745	0.1602	0.3184	0.5276	0.7323	0.88[illegible]2	0.9566	0.9898
ρ = -0.85	3:3	0.1133	0.3488	0.6825	0.9120	0.9877	0.9990	1	1
	3:2	0.1141	0.3496	0.6848	0.9135	0.9882	0.9991	1	1
	3:1	0.1167	0.3542	0.6899	0.9181	0.9889	0.9992	1	1
	3:0	0.1205	0.3762	0.7226	0.9332	0.9923	0.9997	1	1

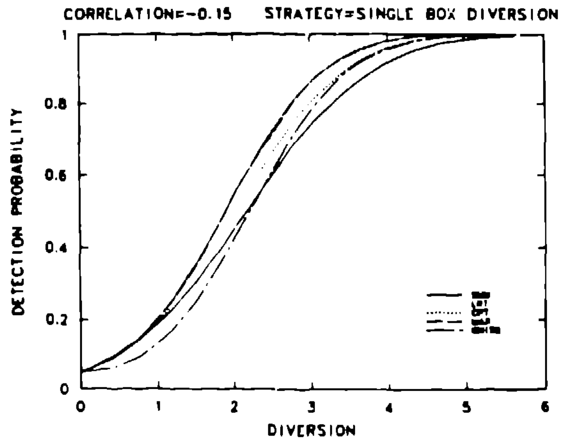

Fig. 2. Detection probabilities for single-box diversion. Units of diversion are in multiples of the common standard deviation.

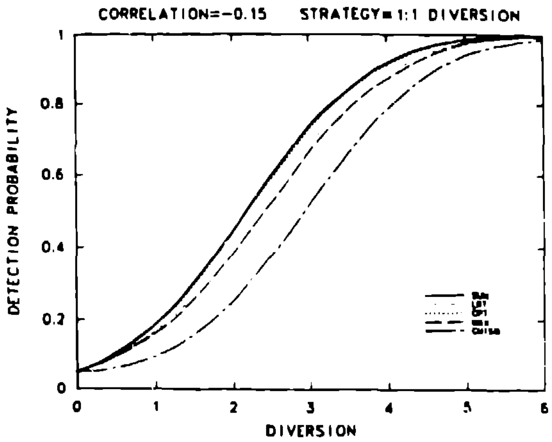

Fig. 3. Detection probabilities for 1:1 diversion. Units of diversion are in multiples of the common standard deviation.

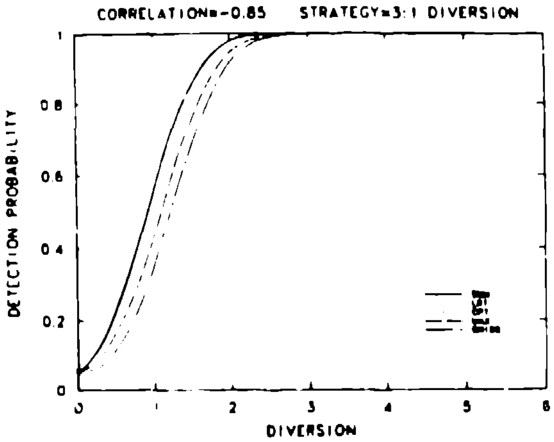

Fig. 4. Detection probabilities for 3:1 diversion. Units of diversion are in multiples of the common standard deviation.